

OIPE

#10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/486,334

DATE: 08/20/2001

TIME: 14:04:56

Input Set : N:\i486334.raw

Output Set: N:\CRF3\08202001\I486334.raw

ENTERED

1 <110> APPLICANT: Droux, Michel
 2 DeRose, Richard
 3 Job, Dominique
 4 <120> TITLE OF INVENTION: Method for Increasing the Content of Sulphur Compounds
 5 and In Particular of Cysteine, Methionine and
 6 Glutathione in Plants and Plants Obtained
 7 <130> FILE REFERENCE: 5500*42
 8 <140> CURRENT APPLICATION NUMBER: US/09/486,334
 9 <141> CURRENT FILING DATE: 2000-07-18
 10 <160> NUMBER OF SEQ ID NOS: 23
 11 <170> SOFTWARE: PatentIn Ver. 2.0
 13 <210> SEQ ID NO: 1
 14 <211> LENGTH: 984
 15 <212> TYPE: DNA
 16 <213> ORGANISM: Arabidopsis thaliana
 17 <220> FEATURE:
 18 <221> NAME/KEY: CDS
 19 <222> LOCATION: (31)..(972)
 20 <400> SEQUENCE: 1
 21 gagagaggat cctctttcca atcataaacc atg gca aca tgc ata gac aca tgc 54
 22 Met Ala Thr Cys Ile Asp Thr Cys
 23 1 5
 24 cga acc ggt aat acc caa gac gat gat tcc cgg ttc tgt tgc atc aag 102
 25 Arg Thr Gly Asn Thr Gln Asp Asp Asp Ser Arg Phe Cys Cys Ile Lys
 26 10 15 20
 27 aat ttc ttt cga ccc ggt ttc tct gta aac cgg aag att cac cac acc 150
 28 Asn Phe Phe Arg Pro Gly Phe Ser Val Asn Arg Lys Ile His His Thr
 29 25 30 35 40
 30 caa atc gaa gat gac gat gat gtc tgg atc aag atg ctt gaa gaa gcc 198
 31 Gln Ile Glu Asp Asp Asp Val Trp Ile Lys Met Leu Glu Glu Ala
 32 45 50 55
 33 aaa tcc gat gtt aaa caa gaa ccc att tta tca aac tac tac tac gct 246
 34 Lys Ser Asp Val Lys Gln Glu Pro Ile Leu Ser Asn Tyr Tyr Tyr Ala
 35 60 65 70
 36 tcg atc aca tct cat cga tct tta gag tct gct tta gct cac atc ctc 294
 37 Ser Ile Thr Ser His Arg Ser Leu Glu Ser Ala Leu Ala His Ile Leu
 38 75 80 85
 39 tcc gta aag ctc agc aat tta aac cta cca agc aac aca ctc ttc gaa 342
 40 Ser Val Lys Leu Ser Asn Leu Asn Leu Pro Ser Asn Thr Leu Phe Glu
 41 90 95 100
 42 ctg ttc ata agc gtt tta gaa gaa agc cct gag atc atc gaa tcc acg 390
 43 Leu Phe Ile Ser Val Leu Glu Glu Ser Pro Glu Ile Ile Glu Ser Thr
 44 105 110 115 120
 45 aag caa gat ctt ata gca gtc aaa gaa aga gac cca gct tgt ata agc 438
 46 Lys Gln Asp Leu Ile Ala Val Lys Glu Arg Asp Pro Ala Cys Ile Ser
 47 125 130 135
 48 tac gtt cat tgc ttc ttg ggc ttc aaa ggc ttc ctc gct tgt caa gct 486

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/486,334

DATE: 08/20/2001

TIME: 14:04:56

Input Set : N:\i486334.raw

Output Set: N:\CRF3\08202001\I486334.raw

```

49   Tyr Val His Cys Phe Leu Gly Phe Lys Gly Phe Leu Ala Cys Gln Ala
50           140                      145                      150
51   cat cga ata gct cat acc ctc tgg aaa cag aac aga aaa atc gta gct   534
52   His Arg Ile Ala His Thr Leu Trp Lys Gln Asn Arg Lys Ile Val Ala
53           155                      160                      165
54   tta ttg atc caa aac aga gta tca gaa tct ttc gcc gtc gat att cat   582
55   Leu Leu Ile Gln Asn Arg Val Ser Glu Ser Phe Ala Val Asp Ile His
56           170                      175                      180
57   ccc gga gcg aag atc gga aaa ggg att ctt tta gac cat gcg acg ggc   630
58   Pro Gly Ala Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly
59   185                      190                      195                      200
60   gtg gtg atc gga gag acg gcg gtg gtt gga gac aat gtt tcg att cta   678
61   Val Val Ile Gly Glu Thr Ala Val Val Gly Asp Asn Val Ser Ile Leu
62           205                      210                      215
63   cac gga gtg acc ttg gga gga aca ggg aaa cag agt ggt gat cgg cat   726
64   His Gly Val Thr Leu Gly Gly Thr Gly Lys Gln Ser Gly Asp Arg His
65           220                      225                      230
66   ccg aag att ggt gat ggt gtg ttg att gga gct ggg agt tgt ata ttg   774
67   Pro Lys Ile Gly Asp Gly Val Leu Ile Gly Ala Gly Ser Cys Ile Leu
68           235                      240                      245
69   ggg aat ata aca atc ggt gag gga gct aag att gga tca ggg tcg gtg   822
70   Gly Asn Ile Thr Ile Gly Glu Gly Ala Lys Ile Gly Ser Gly Ser Val
71           250                      255                      260
72   gtg gtt aag gat gtg ccg gcg cgt acg acg gcg gtt gga aat ccg gcg   870
73   Val Val Lys Asp Val Pro Ala Arg Thr Thr Ala Val Gly Asn Pro Ala
74   265                      270                      275                      280
75   agg ttg att ggt ggg aaa gag aat ccg aga aaa cat gat aag att cct   918
76   Arg Leu Ile Gly Gly Lys Glu Asn Pro Arg Lys His Asp Lys Ile Pro
77           285                      290                      295
78   tgt ctg act atg gac cag aca tcg tat tta acc gag tgg tct gat tat   966
79   Cys Leu Thr Met Asp Gln Thr Ser Tyr Leu Thr Glu Trp Ser Asp Tyr
80           300                      305                      310
81   gtg att taacacaaat gt   984
82   Val Ile
84 <210> SEQ ID NO: 2
85 <211> LENGTH: 314
86 <212> TYPE: PRT
87 <213> ORGANISM: Arabidopsis thaliana
88 <400> SEQUENCE: 2
89   Met Ala Thr Cys Ile Asp Thr Cys Arg Thr Gly Asn Thr Gln Asp Asp
90           1           5           10           15
91   Asp Ser Arg Phe Cys Cys Ile Lys Asn Phe Phe Arg Pro Gly Phe Ser
92           20           25           30
93   Val Asn Arg Lys Ile His His Thr Gln Ile Glu Asp Asp Asp Val
94           35           40           45
95   Trp Ile Lys Met Leu Glu Glu Ala Lys Ser Asp Val Lys Gln Glu Pro
96           50           55           60
97   Ile Leu Ser Asn Tyr Tyr Tyr Ala Ser Ile Thr Ser His Arg Ser Leu
98           65           70           75           80

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/486,334

DATE: 08/20/2001

TIME: 14:04:56

Input Set : N:\i486334.raw

Output Set: N:\CRF3\08202001\I486334.raw

```

99      Glu Ser Ala Leu Ala His Ile Leu Ser Val Lys Leu Ser Asn Leu Asn
100      85                      90                      95
101      Leu Pro Ser Asn Thr Leu Phe Glu Leu Phe Ile Ser Val Leu Glu Glu
102      100                      105                      110
103      Ser Pro Glu Ile Ile Glu Ser Thr Lys Gln Asp Leu Ile Ala Val Lys
104      115                      120                      125
105      Glu Arg Asp Pro Ala Cys Ile Ser Tyr Val His Cys Phe Leu Gly Phe
106      130                      135                      140
107      Lys Gly Phe Leu Ala Cys Gln Ala His Arg Ile Ala His Thr Leu Trp
108      145                      150                      155                      160
109      Lys Gln Asn Arg Lys Ile Val Ala Leu Leu Ile Gln Asn Arg Val Ser
110      165                      170                      175
111      Glu Ser Phe Ala Val Asp Ile His Pro Gly Ala Lys Ile Gly Lys Gly
112      180                      185                      190
113      Ile Leu Leu Asp His Ala Thr Gly Val Val Ile Gly Glu Thr Ala Val
114      195                      200                      205
115      Val Gly Asp Asn Val Ser Ile Leu His Gly Val Thr Leu Gly Gly Thr
116      210                      215                      220
117      Gly Lys Gln Ser Gly Asp Arg His Pro Lys Ile Gly Asp Gly Val Leu
118      225                      230                      235                      240
119      Ile Gly Ala Gly Ser Cys Ile Leu Gly Asn Ile Thr Ile Gly Glu Gly
120      245                      250                      255
121      Ala Lys Ile Gly Ser Gly Ser Val Val Val Lys Asp Val Pro Ala Arg
122      260                      265                      270
123      Thr Thr Ala Val Gly Asn Pro Ala Arg Leu Ile Gly Gly Lys Glu Asn
124      275                      280                      285
125      Pro Arg Lys His Asp Lys Ile Pro Cys Leu Thr Met Asp Gln Thr Ser
126      290                      295                      300
127      Tyr Leu Thr Glu Trp Ser Asp Tyr Val Ile
128      305                      310
130 <210> SEQ ID NO: 3
131 <211> LENGTH: 974
132 <212> TYPE: DNA
133 <213> ORGANISM: Arabidopsis thaliana
134 <220> FEATURE:
135 <221> NAME/KEY: CDS
136 <222> LOCATION: (31)..(966)
137 <400> SEQUENCE: 3
138      gagagaggat cctcttatcg ccgcgttaat atg cca ccg gcc gga gaa ctc cga 54
139      Met Pro Pro Ala Gly Glu Leu Arg
140      1                      5
141      cat caa tct cca tca aag gag aaa cta tct tcc gtt acc caa tcc gat 102
142      His Gln Ser Pro Ser Lys Glu Lys Leu Ser Ser Val Thr Gln Ser Asp
143      10                      15                      20
144      gaa gca gaa gca gcg tca gca gcg ata tct gcg gca gct gca gat gcg 150
145      Glu Ala Glu Ala Ala Ser Ala Ala Ile Ser Ala Ala Ala Ala Asp Ala
146      25                      30                      35                      40
147      gaa gct gcc gga tta tgg aca cag atc aag gcg gaa gct cgc cgt gat 198
148      Glu Ala Ala Gly Leu Trp Thr Gln Ile Lys Ala Glu Ala Arg Arg Asp

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/486,334

DATE: 08/20/2001

TIME: 14:04:56

Input Set : N:\i486334.raw

Output Set: N:\CRF3\08202001\I486334.raw

| | | | | | | | | |
|-----|---|-----|--|-----|--|-----|--|-----|
| 149 | | 45 | | 50 | | 55 | | |
| 150 | gct gag gcg gag cca gct tta gct agc tat cta tat tcg acg att ctt | | | | | | | 246 |
| 151 | Ala Glu Ala Glu Pro Ala Leu Ala Ser Tyr Leu Tyr Ser Thr Ile Leu | | | | | | | |
| 152 | | 60 | | 65 | | 70 | | |
| 153 | tct cat tcg tct ctt gaa cga tct atc tcg ttt cat cta gga aac aag | | | | | | | 294 |
| 154 | Ser His Ser Ser Leu Glu Arg Ser Ile Ser Phe His Leu Gly Asn Lys | | | | | | | |
| 155 | | 75 | | 80 | | 85 | | |
| 156 | ctt tgt tcc tca acg ctt tta tcc aca ctt tta tac gat ctg ttc tta | | | | | | | 342 |
| 157 | Leu Cys Ser Ser Thr Leu Leu Ser Thr Leu Leu Tyr Asp Leu Phe Leu | | | | | | | |
| 158 | | 90 | | 95 | | 100 | | |
| 159 | aac act ttt tcc tcc gat cct tct ctt cgt aac gcc acc gtc gca gat | | | | | | | 390 |
| 160 | Asn Thr Phe Ser Ser Asp Pro Ser Leu Arg Asn Ala Thr Val Ala Asp | | | | | | | |
| 161 | | 105 | | 110 | | 115 | | 120 |
| 162 | cta cgc gct gct cgt gtt cgt gat cct gct tgt atc tcg ttc tct cat | | | | | | | 438 |
| 163 | Leu Arg Ala Ala Arg Val Arg Asp Pro Ala Cys Ile Ser Phe Ser His | | | | | | | |
| 164 | | 125 | | 130 | | 135 | | |
| 165 | tgt ctc ctc aat tac aaa ggc ttc tta gct att cag gcg cat cgt gta | | | | | | | 486 |
| 166 | Cys Leu Leu Asn Tyr Lys Gly Phe Leu Ala Ile Gln Ala His Arg Val | | | | | | | |
| 167 | | 140 | | 145 | | 150 | | |
| 168 | tca cac aag cta tgg aca caa tca cgg aag cca tta gca tta gct cta | | | | | | | 534 |
| 169 | Ser His Lys Leu Trp Thr Gln Ser Arg Lys Pro Leu Ala Leu Ala Leu | | | | | | | |
| 170 | | 155 | | 160 | | 165 | | |
| 171 | cac tca aga atc tcc gat gta ttc gct gtt gat atc cat cca gca gcg | | | | | | | 582 |
| 172 | His Ser Arg Ile Ser Asp Val Phe Ala Val Asp Ile His Pro Ala Ala | | | | | | | |
| 173 | | 170 | | 175 | | 180 | | |
| 174 | aag atc gga aaa ggg ata ctt cta gac cac gca acc gga gtt gta gtc | | | | | | | 630 |
| 175 | Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly Val Val Val | | | | | | | |
| 176 | | 185 | | 190 | | 195 | | 200 |
| 177 | gga gaa aca gcg gtg att ggg aac aat gtt tca atc ctt cac cat gtg | | | | | | | 678 |
| 178 | Gly Glu Thr Ala Val Ile Gly Asn Asn Val Ser Ile Leu His His Val | | | | | | | |
| 179 | | 205 | | 210 | | 215 | | |
| 180 | aca cta ggt gga aca ggt aaa gct tgt gga gat aga cat ccg aag atc | | | | | | | 726 |
| 181 | Thr Leu Gly Gly Thr Gly Lys Ala Cys Gly Asp Arg His Pro Lys Ile | | | | | | | |
| 182 | | 220 | | 225 | | 230 | | |
| 183 | ggt gac ggt tgt ttg att gga gct gga gcg act att ctt gga aat gtg | | | | | | | 774 |
| 184 | Gly Asp Gly Cys Leu Ile Gly Ala Gly Ala Thr Ile Leu Gly Asn Val | | | | | | | |
| 185 | | 235 | | 240 | | 245 | | |
| 186 | aag att ggt gca ggt gct aaa gta gga gct ggt tct gtt gtg ctg att | | | | | | | 822 |
| 187 | Lys Ile Gly Ala Gly Ala Lys Val Gly Ala Gly Ser Val Val Leu Ile | | | | | | | |
| 188 | | 250 | | 255 | | 260 | | |
| 189 | gac gtg cct tgt cga ggt act gcg gtt ggg aat ccg gcg aga ctt gtc | | | | | | | 870 |
| 190 | Asp Val Pro Cys Arg Gly Thr Ala Val Gly Asn Pro Ala Arg Leu Val | | | | | | | |
| 191 | | 265 | | 270 | | 275 | | 280 |
| 192 | gga ggg aaa gag aag cca acg att cat gat gag gaa tgt cct gga gaa | | | | | | | 918 |
| 193 | Gly Gly Lys Glu Lys Pro Thr Ile His Asp Glu Glu Cys Pro Gly Glu | | | | | | | |
| 194 | | 285 | | 290 | | 295 | | |
| 195 | tcg atg gat cat act tca ttc atc tcg gaa tgg tca gat tac atc ata | | | | | | | 966 |
| 196 | Ser Met Asp His Thr Ser Phe Ile Ser Glu Trp Ser Asp Tyr Ile Ile | | | | | | | |
| 197 | | 300 | | 305 | | 310 | | |

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/486,334

DATE: 08/20/2001

TIME: 14:04:56

Input Set : N:\i486334.raw

Output Set: N:\CRF3\08202001\I486334.raw

974

```

198      taaagttg
200 <210> SEQ ID NO: 4
201 <211> LENGTH: 312
202 <212> TYPE: PRT
203 <213> ORGANISM: Arabidopsis thaliana
204 <400> SEQUENCE: 4
205      Met Pro Pro Ala Gly Glu Leu Arg His Gln Ser Pro Ser Lys Glu Lys
206      1          5          10          15
207      Leu Ser Ser Val Thr Gln Ser Asp Glu Ala Glu Ala Ala Ser Ala Ala
208      20          25          30
209      Ile Ser Ala Ala Ala Ala Asp Ala Glu Ala Ala Gly Leu Trp Thr Gln
210      35          40          45
211      Ile Lys Ala Glu Ala Arg Arg Asp Ala Glu Ala Glu Pro Ala Leu Ala
212      50          55          60
213      Ser Tyr Leu Tyr Ser Thr Ile Leu Ser His Ser Ser Leu Glu Arg Ser
214      65          70          75          80
215      Ile Ser Phe His Leu Gly Asn Lys Leu Cys Ser Ser Thr Leu Leu Ser
216      85          90          95
217      Thr Leu Leu Tyr Asp Leu Phe Leu Asn Thr Phe Ser Ser Asp Pro Ser
218      100         105         110
219      Leu Arg Asn Ala Thr Val Ala Asp Leu Arg Ala Ala Arg Val Arg Asp
220      115         120         125
221      Pro Ala Cys Ile Ser Phe Ser His Cys Leu Leu Asn Tyr Lys Gly Phe
222      130         135         140
223      Leu Ala Ile Gln Ala His Arg Val Ser His Lys Leu Trp Thr Gln Ser
224      145         150         155         160
225      Arg Lys Pro Leu Ala Leu Ala Leu His Ser Arg Ile Ser Asp Val Phe
226      165         170         175
227      Ala Val Asp Ile His Pro Ala Ala Lys Ile Gly Lys Gly Ile Leu Leu
228      180         185         190
229      Asp His Ala Thr Gly Val Val Val Gly Glu Thr Ala Val Ile Gly Asn
230      195         200         205
231      Asn Val Ser Ile Leu His His Val Thr Leu Gly Gly Thr Gly Lys Ala
232      210         215         220
233      Cys Gly Asp Arg His Pro Lys Ile Gly Asp Gly Cys Leu Ile Gly Ala
234      225         230         235         240
235      Gly Ala Thr Ile Leu Gly Asn Val Lys Ile Gly Ala Gly Ala Lys Val
236      245         250         255
237      Gly Ala Gly Ser Val Val Leu Ile Asp Val Pro Cys Arg Gly Thr Ala
238      260         265         270
239      Val Gly Asn Pro Ala Arg Leu Val Gly Gly Lys Glu Lys Pro Thr Ile
240      275         280         285
241      His Asp Glu Glu Cys Pro Gly Glu Ser Met Asp His Thr Ser Phe Ile
242      290         295         300
243      Ser Glu Trp Ser Asp Tyr Ile Ile
244      305         310
246 <210> SEQ ID NO: 5
247 <211> LENGTH: 1048
248 <212> TYPE: DNA

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/486,334

DATE: 08/20/2001
TIME: 14:04:57

Input Set : N:\i486334.raw

Output Set: N:\CRF3\08202001\I486334.raw

L:590 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
L:710 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11